UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P O Box 1450
Alexandria, Virginia 22313-1450
www uspto gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/647,586	08/25/2003	Michael Seltzer	M61.12-0550	2416

27366 7590 06/17/2008
WESTMAN CHAMPLIN (MICROSOFT CORPORATION)
SUITE 1400
900 SECOND AVENUE SOUTH
MINNEAPOLIS, MN 55402-3244

EXAMINER	
SHAH, PARAS D	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
06/17/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/647,586	SELTZER ET AL.	
	Examiner	Art Unit	
	PARAS SHAH	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *11 March 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-3,5-13,15-17,20,21 and 23-25* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-3,5-13,15-17,20,21 and 23-25* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

1. This Office Action is in response to the Amendments and Arguments filed on 03/11/2008. Claims 1-3, 5-13, 15-17, 20, 24, and 25 remain pending, while claims 21 and 23 have been cancelled. The Applicants' amendment and remarks have been carefully considered, but they do not place the claims in condition for allowance. Accordingly, this action has been made FINAL.

2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

Response to Arguments

3. Applicant's arguments (pages 7-12) filed on 03/11/2008 with regard to claims 1-3, 5-13, 15-17, 20, 24, and 25 have been fully considered but they are moot in view of new grounds for rejection.

Response to Amendment

4. Applicants' amendments filed on 03/11/2008 have been fully considered. The newly amended limitations in claims 1 and 13 are moot in view of new grounds of rejection. Specifically, the limitations, "for each frame, " "fixed scaling parameter for the random component...wherein the fixed scaling parameter is the same for all frames and is les than one."

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1-3, 5, 6, 11, 12, 15, 16, 20, and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over by Laroche *et al.* ("HNM: A Simple Efficient Harmonic and Noise Model for Speech" 1993) in view of Rao Gadde *et al.* (US 7,120,580) in view of Gao (US 2002/0035470) and in view of Rezayee ("An Adaptive KLT Approach for Speech Enhancement").

As to claims 1 and 13, Laroche *et al.* teaches a method of identifying an estimate for a noise reduced value representing a noise-reduced speech signal, the method comprising:

decomposing each frame (see page 1, equation 1, and 2, each signal represents a particular value at each time t. It would have been obvious to take a frame of a signal to capture the pattern of series of speech samples.) of a noisy speech signal (see page 1, left column, sect. 1, line 2-3) into a harmonic component (see page 1, left column, sect. 1, line 2) and a random component (see page 1, left column, sect. 1, line 2-3) (e.g. It should be noted that noise contained in speech is non-periodic and is random);

for each frame, summing the harmonic component and the random component to form the noise-reduced value (see page 3, right column, sect. 4,

lines 6-8) (e.g. The synthetic signal is formed by the harmonic component and the random component from equation 1 and 2. Equations 1 and two could have been modified to take into account frame based summing rather then per sample.).

However, Laroche does not specifically teach the determination of the scaling parameter, and multiplication of the harmonic and random component.

Rao Gadde *et al.* does teach determining a scaling parameter (see Figure 4, weight (W) and see col. 4, lines 15-20, SNR is used to compute the weights) for at least the harmonic component (see Figure 4, model weight (e.g. The model weight is used for the speech portion (see Abstract));

multiplying the harmonic component by the scaling parameter for the harmonic component to form a scaled harmonic component (see Figure 4, multiplier 430 multiplies model weight to speech model 320);

multiplying the random component by a scaling parameter for the random component to form a scaled random component (see Figure 4, multiplier 430 multiplies noise weight to the noise model 410) (e.g. The two scaling parameters or weights for the speech and noise are different, where the noise weight is 1-W and the speech weight is W);

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the identification of an estimate for a noise-reduced speech signal as taught by Laroche with the use of scaling parameters for the harmonic and random component as taught by Rao Gadde *et*

al. The motivation to have included these scaling parameters involves the ability to adjust the noise models based on SNR at low energy of the input signal for speech proper recognition of speech, depending on noise types (see col. 4, lines 5-21 and col. 1, lines 25-53).

However, Laroche in view of Rao Gadde *et al.* do not specifically teach determining a scaling parameter for each frame and the scaling parameter for the random component being fixed and less than one.

Gao does teach the determining a scaling parameter for each frame (see [0053], speech detected for each frame and a corresponding gain computed) and the scaling parameter for the random component being fixed and less than one (see [0053] and [0054]) (e.g. The scaling factor is fixed since the NSR is 1, for noise frames (random) and C is preselected and stored, where C is 0.4-0.6. Using the equation for the gain. Gf=1-(0.4:0.6), which is less then 1).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the estimation of a noise reduced speech signal presented by Laroche in view of Rao Gadde *et al.* with the reduction of noise relative to the noisy speech signal as taught by Rezayee. The motivation to have combined the two references involves the improvement in the quality of the voice signal (see page 2, [0021]).

However, Laroche in view of Rao Gadde *et al.* in view of Gao *et al.* do not specifically teach the harmonic scaling parameter being a ratio of an energy of

the harmonic component without the random component of the frame to an energy of the noisy speech signal.

Rezayee does teach the harmonic scaling parameter being a ratio of an energy of the harmonic component without the random component of the frame to an energy of the noisy speech signal (see Table II, equation for gi(n)) (e.g. For each sample, the gain is calculated using the energy of the speech divided by the total signal.).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the estimation of a noise reduced speech signal presented by Laroche in view of Rao Gadde *et al.* in view of Gao with the substitution of the scaling gain as taught by Rezayee. The motivation to have combined the references involves the improvement in the quality of the voice signal (see Gao page 2, [0021]).

As to claim 2, Laroche *et al*. in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claim 1, above.

Furthermore, Laroche teaches modeling the harmonic component as a sum of harmonic sinusoids (see page 1, right column, sect. 1, lines 1-3 and equation 1) (e.g. It is apparent that equation 1 can be put in terms of cosine and sine using Euler's Relation. Also, the multiplication of the harmonic component by the parameter is done through the multiplication of the individual harmonic sinusoids. Further, the multiplication of the parameters before the summing or

after the summing is equivalent since the A_k could have been factored out and multiplied later as a vector.)

As to claims 3 and 15, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claims 1 and 13, above.

Furthermore, Laroche teaches determining a least-squares solution to identify the harmonic component (see page 2, left column, sect. 3, 2nd paragraph, lines 1-5) (e.g. It should be noted that a least squares method is used to estimate the parameters to obtain the harmonic component. The voiced segment is the harmonic component).

As to claims 5 and 20, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claims 1 and 13, above.

Furthermore, Rezayee teaches a gain factor determined from a the ratio of the harmonic component to the energy of the noisy speech signal (see Table II, equation for gi(n)) (e.g. For each sample, the gain is calculated using the energy of the speech divided by the total signal. It would have been obvious to one of ordinary skilled in the art to sum up the energy values of the samples to get a more robust estimate for the gain for a frame of a signal or a specific window of time.)

As to claims 6 and 16, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claims 1 sand 13, above.

Furthermore, Laroche teaches decomposing a vector of time samples from a frame (see page 1, right column, sect. 2, line 5-8) of the noisy speech signal into a harmonic component vector of time samples and a random component vector of time samples (e.g. The use of time samples is used as seen by the summation bounds for equation 1. Further, the random signal is obtained from the subtraction of the original speech signal by the harmonic part, which is also a specific time sample size).

As to claims 11 and 24, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claims 1 and 13, above.

Furthermore, Laroche teaches is used to perform speech recognition (see page 3, right column, sect. 6, line, 1st paragraph, line 2) (e.g. The use of speech enhancement directly relates to speech recognition). The feature vector referred to is described as being the signal representing the noise reduced signal, which is obtained from the steps stated in claim 1.

As to claim 12, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claim 1, above.

Furthermore, Laroche teaches using the noise-reduced value (e.g. also known as the synthesized signal described above) in speech coding (see page 3,

right column, sect. 6, line, 1st paragraph, line 2) (e.g. It is apparent that the use of the HNM model for speech enhancement and timbre modification directly relates to speech coding since the intelligibility of the signal is of importance, which relates to the noise component).

7. Claims 7-10, 17, and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee as applied to claims 6, 16, and 24 above, and further in view of Seltzer (CMU Speech Group 1999).

As to claims 7 and 17, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee in view of Seltzer teach all of the limitations as in claims 6 and 16, above.

Furthermore, Seltzer does teach the determination of the Mel spectrum from a speech signal (see page 1,sect. 2, Block Diagram).

It would have been obvious to one of ordinary skilled in the art to have modified the harmonic plus noise model as taught by Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee by the determination of the Mel spectrum as taught by Seltzer to find the Mel Spectrum. The motivation to combine the two involves the extraction of features from the speech signal. The Mel spectrum also enables the detection of voiced segments allows the frequency amplitudes to be seen. The speech signal in this case is the synthesized harmonic component. This method is commonly used in speech recognition systems.

As to claim 8, Laroche *et al.* in view of Rao Gadde in view of Gao in view of Rezayee teach all of the limitations as in claims 7 and 17, above

However, Laroche does not specifically teach the multiplication of the Mel Spectrum for the harmonic component with the scaling factor.

Seltzer does teach the calculation of the Mel Spectrum (see Page 2, sect. 4d., see equation) by the harmonic component with the scaling factor pre-multiplied from an input speech signal. The speech signal in this case is the harmonic component that was modeled.

The multiplication of the scaling factor could have been pre-multiplied as the frequency content of the signal will not change, but rather the amplitude. However, since the scaling factor applies to all frequency components, the scaling factor can be also multiplied after the Mel Spectrum is obtained, which will allow the same result to be obtained. The multiplying of the scaling parameter to the harmonics or the multiplication of the scaling parameter after the Mel coefficients is found from the harmonics is equivalent. The scaling parameter is the A_k value shown in claim 1. Further, the motivation to have combined the two references was stated in the rejection for claim 7 for recognizing the synthesized speech created from claim 1.

As to claims 9,10 and 25, Laroche *et al*. in view of Rao Gadde in view of Gao in view of Rezayee in view of Seltzer teach all of the limitations as in claims 6 and 16, above

Furthermore, Seltzer teaches the forming of Mel Frequency Cepstral Coefficients feature vector (see Page 3, sect. 4e, equation) from a speech signal for speech recognition (page 1, sect. 1, line 1). This is found from the Mel Spectrum.

Conclusion

8. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL.** See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PARAS SHAH whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:00a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571)272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Paras Shah/
Examiner, Art Unit 2626

06/13/2008

/Patrick N. Edouard/
Supervisory Patent Examiner, Art Unit 2626